DOCUMENT RESUME

ED 202 595 PS 012 197

AUTHOR Gersten, Russell M.; And Others
TITLE The Relationship of Entry IQ Level and Yearly Academic Growth Rates of Children in a Direct Instruction Model. A Longitudinal Study of Over 1500 Children.
PUB DATE Apr 81
NOTE 68p.; Paper presented at the Annual Meeting of the American Educational Research Association (Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Achievement Gains; Cognitive Development; *Elementary School Mathematics; *Elementary School Students; Individual Differences; *Intelligence Quotient; *Kindergarten Children; Longitudinal Studies; *Low Income Groups; Primary Education; *Reading Ability; Reading Improvement; Standardized Tests
IDENTIFIERS *Direct Instruction Model; Metropolitan Achievement Tests; Slosson Intelligence Test; Wide Range Achievement Test

ABSTRACT

This study examines the relationship between IQ and yearly academic growth rate in reading and mathematics for low income children in the primary grades involved in the Direct Instruction Follow Through program (DIFT). Low income children, from 20 communities in the United States, who engaged in the DIFT program between 1969 and 1977 for either 3 years (grades 1 through 3) or 4 years (kindergarten through grade 3), participated in the study. Upon entering the program, in either kindergarten or first grade, students were tested on the Wide Range Achievement Test (WRAT) and the Slosson Intelligence Test (SIT). Each spring students were tested on the WRAT, SIT, and (beginning in grade 1) the Metropolitan Achievement Test (MAT). For the purposes of the longitudinal analyses of variance, children's entry IQ score was used to classify them into one of six IQ blocks (70 and lower, 71-90, 91-100, 111-130, 131 and above), and then a mixed (split-plot) analysis of variance design was used to analyze results with one between-groups factor (IQ block) and one within-groups factor (time of test). The major finding of this study was that little, if any, relationship existed between entry IQ and yearly learning rate. Generally, the students entering the program at higher cognitive skill levels finished third grade at higher levels, but the growth rates (or amounts of materials learned) were not significantly different between IQ blocks. (Author/MP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED202595

The Relationship of Entry IQ Level and Yearly Academic Growth Rates of Children in a Direct Instruction Model: A Longitudinal Study of Over 1500 Children

Russell M. Gersten
Thomas J. Heiry[1]
Wesley C. Becker
W. A. T. White
University of Oregon

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Russell M. Gersten
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

PS012197

Paper presented at annual conference of AERA: Los Angeles, CA, April 1981

[1]Dallas Independent Schools

Over forty years ago, Anderson (1939) reported that though mental age was correlated with general level of achievement, it was not strongly related to year-to-year gains in learning. He asserted that those who begin school at a higher skill level tend to end at higher levels, but do not necessarily gain more during each year of school. His hypothesis has rarely been tested in a serious way. As Jensen (1980, p. 472) says, "The published evidence here is surprisingly meager..." One of the deterrents to this type of study is the manifold technical problems in the use of gain scores to assess academic growth (Becker & Engelmann, 1976; Cronbach & Furby, 1970; Linn & Slinde, 1977; Bereiter, 1973).

The present study examines the relationship between entry IQ and yearly achievement gains for samples of low income primary grade students who were enrolled in a highly structured educational program. Rather than use gain scores, longitudinal analyses of variance were performed on achievement testing as within-group factors. Students were placed in blocks on the basis of entry IQ scores (e.g., less than 71, 71-90, etc.). Our hypothesis, following Anderson's (1939) theory, was that the yearly academic growth rate would be roughly equivalent for all IQ blocks. The low IQ students would begin at a lower academic level, and quite likely complete the third grade at a lower academic level than their peers, but--with direct instruction in small ability groups and mastery learning--year-to-year gains would be comparable for all IQ blocks. However, rather than use gain scores, the achievement was analyzed by using a mixed two-way analysis of variance with one between group factor (IQ block) and one within group factor. Thus, in statistical terms, our hypotheses became: there would be no significant IQ block X time-of-testing interactions for any of the three achievement tests, although there would undoubtedly be significant main effects for the IQ block

factor alone and the time-of-testing factor alone. We suspect that the relationships found here might not hold for similar students in different types of educational programs.

Relevant Research

There is a reasonable body of evidence showing strong correlations between IQ and current achievement scores; often the strength of the correlations is in the high to moderate range for both Anglo and non-Anglo children. Generally, concurrent validity coefficients range from .6 to .8, with slightly higher coefficients for Anglo students than Black students (Vane, 1966; Thorndike, 1971b; Crano, Kenny & Campbell, 1972). For example, Bossard, Reynolds, and Gutkin (1980) report concurrent validity coefficients between the Stanford-Binet and the Reading and Arithmetic subtests on the Wide Range Achievement Test of .74 and .70 for Black students referred for special educational services in the elementary grades and .82 and .82 for comparable Anglo students. Yet, none of these studies deals with the issues of prediction of yearly learning in schools.

Several researchers have attempted to explore the relationship between either IQ or mental age and acquisition of new skills, but always in laboratory settings. There appears to be a strong relationship for moderately and severely retarded individuals (Zeaman, 1973; Wilhelm & Lovaas, 1976; Gersten, 1980). The evidence for non-handicapped children, on the other hand, is quite variable (Stevenson & Odom, 1965; Gulliksen, 1968; Olson et al., 1968; Carrol, 1980). Summarizing this research, Cronbach (1976) stated, "We conclude that 'learning rate' or 'ability to learn' is not a satisfactory construct. Learning of what? and from what instruction?"

The latter questions become particularly salient when one examines the two studies that actually explored the predictive validity of IQ for elementary

school children in actual school settings. Crano et al. (1972) looked at both predictive and concurrent validities between Total IQ on the Lorge-Thorndike and achievement scores on the Iowa Test of Basic Skills. Concurrent validity correlations between IQ in Grade 4 and total achievement scores on the Iowa tests were .73 for a sample of primarily Anglo suburban students (N=3992), and a moderately strong .62 for a primarily Black inner city sample (N=1449). Predictive validity coefficients between fourth grade IQ and sixth grade achievement were as high as concurrent validity indices; r = .75 for the entire sample. When Crano et al. (1972) performed a subtest by subtest breakdown, they found that particular predictive validity coefficients tended to be slightly higher for Blacks than Anglos. But the study lacks any information on the quality of instruction the students were receiving in grades 5 and 6. Cronbach (19 76) observes:

> There is a suspicion afoot that education calls for analytic ability just because the materials are capable of being put into meaningful relationships and the instructor has either failed to so display the relationships or has given an explanation that is hard to follow. Then one has to use his (sic) brains!

Oakland's (1978) study of minority students in the primary grades does not support the contentions of Crano et al. (1972) about predictive validity of IQ with intermediate grade students. He looked at predictive validity between IQ in kindergarten (on the Slosson Intelligence Test) and second grade and fourth grade achievement. Table 1 summarizes the results. Note

---

Table 1 about here

---

that predictive validities are not at all significant for the Black sample. Cautions about generalizing from this study are limited by

the relatively small sample sizes and the failure to specify the type of instruction or the quality of instruction the students received.

As Anderson (1939) observed, neither concurrent nor predictive validity studies directly confront the issues of relationship between IQ and academic growth. The laboratory studies on learning rate of Zeaman (1973), Meyer (in Becker & Engelmann, 1976); Olson et al. (1968), and Stevenson & Odom (1965) offer little on how students learn in applied educational settings. Only the recent provocative learning potential studies by Budoff and his colleagues (Budoff & Corman, 1976; Budoff & Hamilton, 1976) have attempted to look at learning rate in actual classroom situations; and even here, the type of instruction offered in these classrooms is not specified.

Overview of the Study

The present study is an attempt to look at the relationship between IQ and yearly academic growth rate in reading and mathematics for low income children in the primary grades involved in the Direct Instruction Follow Through model. This model is a highly structured educational program utilizing small group instruction with high levels of academic engaged time and the principles of mastery learning to ensure that all children in the program--including those who enter with low cognitive skills or limited language backgrounds--received adequate amounts of daily instruction at the appropriate level. Curriculum materials are based on systematic teaching of general problem solving strategies and are designed to begin at a lower level than a typical basal series. Student progress through the curriculum is determined by the use of criterion-referenced tests administered every six weeks in each academic area. (See Becker & Carnine, 1980 for a more complete description of the program.)

Our research hypothesis was that, with a mastery-based educational model, and with allocation of personnel to insure that all children are taught at an appropriate difficulty level (Fischer et al., 1980), yearly academic gains will be roughly similar for children regardless of their entry ability. Predictably, students entering with higher IQ's would also begin with higher entry scores on the academic tests. However, according to the hypothesis, there would be no significant interaction between entry IQ and yearly academic growth rates as measured by standardized achievement tests.

The analyses in this paper are based on achievement test and intelligence test scores on computer file for the 7000 low income children involved in 20 communities in the Direct Instruction Follow Through Project between 1969 and 1977.

Upon entering the program, either in kindergarten or first grade, students were tested on the Wide Range Achievement Test (WRAT) and the Slosson Intelligence Test (an abridged adaptation of the Stanford-Binet). Each spring students were tested on the WRAT, Slosson, and, beginning in grade 1, the Metropolitan Achievement Test (MAT). (No kindergarten level versions of the MAT were available until 1977.) The WRAT was used to test oral word reading and arithmetic computation; the Metropolitan was used to test in a more comprehensive fashion reading comprehension and a wide range of math skills. The Slosson was used as a test of general language competence. It is perhaps important to state at this point that these tests were used for purposes of evaluation research only; all placement decisions were based on individually administered criterion-referenced tests.

Though our concern was with yearly academic growth, for a variety of well known reasons related to the low (or uncertain) reliability of gain scores (Cronbach & Furby, 1976; Bereiter, 1973; Linn & Slinde, 1977), we did not use gain scores in our analyses. Instead, we looked at data in two different ways:

(a) a series of longitudinal two-way mixed analyses of variance using IQ block at entry as the between group factor, and time of test as the repeated factor,

(b) multiple correlations following the procedures of Crano et al. (1972), looking at concurrent and predictive validities of IQ and achievement tests.

For the first series of analyses, children were placed in one of six IQ blocks (less than 71, 71-90, 91-100, 101-110, 111-130, and above 130) on the basis of entry IQ. Our hypotheses were that there would be no (or very few) significant interactions between entry IQ and yearly achievement scores, and that only a very small proportion of the variance would be explained by the interaction. This hypothesis stems from Anderson's (1939) work and Cronbach's (1976) formulations about intellectual growth in appropriate educational settings. It extends the laboratory work of Stevenson and Odom (1965) and Olson et al. (1968) into applied educational field settings.

In terms of the correlational analyses, our predictions were that, in controlled educational settings concurrent validities would always be higher than predictive validities. This is in opposition to the empirical findings of Crano et al. (1972) for both Anglo and Black students in unspecified settings (who found predictive validities to be lower), though supported in part by Oakland's (1978) findings for small samples of Black

and Hispanic children. Also, following the work of Budoff and Hamilton (1976) on learning potential and appropriate assessment of non-white students, we were interested in comparing the predictive validity of IQ and achievement test scores administered to children upon entry to school versus those administered after one full year of instruction.

Details of testing procedure and schedules and data analysis procedures are presented below.

Method

Sample

The sample consisted of low income students engaged in the Direct Instruction Follow Through program for either three or four years between 1969 and 1977. This includes five cohorts of children. (Students in the same cohort begin the program in the same year.) Twenty communities were involved; twelve offered four year programs (kindergarten through third grade). The remaining eight, located in the south and southwest offered a three-year program, beginning in first grade. Table 2 lists the communities, arranged by primary ethnic compositions.

Table 3 about here

All 20 communities offered the Direct Instruction program; degree of implementation was monitored by consultants from the Follow Through sponsor. As can be seen from the list of sites, the EF sites tend to come from the South and Southwest. The Black communities in the EK sites (4 year programs) are primarily urban; Black students in EF sites are primarily rural. Virtually all the Anglo children, Hispanic children, and Native American children are from rural communities.

Table 3 offers a description of the EF and EK samples. For the analyses

Table 3 about here

to be considered part of the sample, a student was required to

1) qualify as low income by Office of Education Follow Through guidelines
2) attend the full four years of the program (3 years in EF sites)
3) take the Slosson Intelligence Test during the first six weeks of the program.

There were five cohorts of children involved in the study. Cohort I began in 1969; cohort V in 1973. The last cohort (V, EK) completed the program in 1977.

Testing

Each year the testing was guided by a manual of procedures, built around the test publishers' manuals so as to be consistent with all standardization requirements. Training began at the University of Oregon with the field supervisors, who were responsible for testing at the sites. On site, a similar training program occurred for the local staff assisting in the testing. All teachers were trained to give the MAT, and the aides were trained to test the students on the individually administered tests, the SIT and WRAT. The exact testing schedule for EK and EF sites is presented in Table 4. The MAT was given in the first two weeks in April, the norm dates specified by the publisher. The WRAT and SIT were each

Table 4 about here

given to students during the first four weeks of their first school year in Follow Through and then during the last six weeks of every school year in Follow

Through. (The WRAT and SIT are normed on the basis of chronological age, not grade level.)

Table 5 indicates which cohorts were tested on which tests. WRAT and SIT testing began in 1969; the MAT was not added until 1972; budgetary cuts in 1975 precluded some individualized testing. (For details, see Becker & Engelmann, 1978.)

Table 5 about here

Strategies for Data Analysis

For the purposes of the longitudinal analyses of variance, children were classified into one of six IQ blocks based on entry IQ: 70 and lower, 71-90, 91-100, 101-110, 111-130, 131 and above. (The latter IQ block was omitted for EF sites because the sample size was less than 15 and therefore too small to provide reliable estimates.) For each dependent measure--MAT Total Reading, MAT Total Math, WRAT Reading, and Slosson IQ--a mixed (split-plot) analysis of variance design (Winer, 1971) was used to analyze results with one-between groups factor (IQ block) and one within-groups factor (time of test). The BMD computer program (Dixon, 1972) was utilized for these analyses. Only students present for all testings on a particular measure were included in the sample for that measure. Thus, the sample size fluctuates depending on the test. EF and EK samples were always analyzed separately.

In addition to reporting the F and p value for each interaction, the magnitude of effect (proportion of explained variance) was calculated using Friedman's (1964) procedure for use with large sample sizes. Where significant interactions were found, Scheffe post-hoc tests

were performed to determine the locus of the interaction. The Scheffe tests used a weighted sample mean, since cell sizes were clearly unequal; in fact, there was roughly a normal distribution centering around a mean IQ of 100 for the EK sample and 95 for the EF sample. The pooled error term was used in the Scheffe tests, as Winer (1971) suggests for mixed (split-plot) designs.

Correlational Analyses. Again, the EF and EK samples were always analyzed separately. This was of special interest here, since psychometric literature (e.g., McCall, Appelbaum & Hobarty, 1973) suggests that the predictive validity of IQ increases at the age of 6 (the common entry age for EF sites). For the predictive validity/concurrent validity sample, students were required to be present for all testings during third grade and either kindergarten (EK sample) or first grade (EF sample, for all three instruments--the SIT, WRAT, and MAT. However, the sample sizes appeared to be adequate--n = 594 for the EK sample and 735 for the EF sample. Whereas in the ANOVAs we were forced to look at only MAT Total Reading and MAT Total Math (because none of the MAT subtests, e.g., Word Knowledge, appeared for three consecutive years), the more flexible correlational approach allowed for the examination of all MAT subtests--Word Knowledge, Language, Math Concepts, Math Problem Solving, etc.

Results and Discussion

Longitudinal Analyses of Variance

WRAT Reading. Figures 1 and 2 and Table 6 present the results for WRAT Reading, an individually administered test of reading decoding skills. The longitudinal gains and losses in standard score units for each IQ block are plotted in for the EK sample in Figure 1 and for the EF sample in Figure 2.

Table 6, Figures 1 and 2

(The mean for the WRA. tandardization sample is 100; the standard deviation is 15.) Note that the growth patterns are basically parallel for all IQ blocks in both samples. In all cases there is strong growth during the first year with a basic maintenance pattern thereafter.

The ANOVA indicates no significant interaction (i.e., no differential growth rate) for the EK sample; $F_{20,2744} = .82$. For the EF sample, a significant interaction is found; $F_{12,4491} = 4.61$; yet, only one percent of the variance is explained by the interaction. Scheffé tests reveal that the locus of the interaction occurs during first grade; those entering with low IQs (below 70) gain significantly less than those in the highest IQ block (111 and above).

In the area of decoding, the basic pattern one sees is that those entering with lower IQs begin with lower WRAT Reading scores and end with lower WRAT Reading scores, but--with one exception--the yearly gains are not significantly different. This result would not be particularly important if the program was not effective; but there is strong evidence of effectiveness. For these low income children, the overall post-third grade WRAT Reading mean corresponds to the 82nd percentile for the EK sample and 70th percentile for the EF sample, well above the median performance of the standardization sample, though they begin the program well below that level.

Table 7, Figures 3 and 4

MAT Total Math. Figures 3 and 4 and Table 7 present the descriptive statistics and results for MAT Total Math. Certain differences in the conventions used for the MAT need to be explained at this point. No entry (baseline) scores

are available on this measure since the 1970 edition of the MAT did not provide a level of the test appropriate for either the beginning of kindergarten or beginning of first grade. (This has been rectified in the 1977 version of the test.) Thus, Figures 3 and 4 follow the children through only the last two years of the program--second and third grades. The MAT Math subtests cover a broad range of operations--problem solving, geometric and measurement concepts, as well as arithmetic computation.

Note that the lines in Figure 3 for the EK sample are virtually all parallel. The mean growth for all IQ blocks is roughly the same for both second and third grade; when the mean standard scores are converted to grade-equivalents, this corresponds to a mean growth rate of 1.0 grade equivalents per year for all IQ blocks, even those with IQs below 70. Again, no significant interaction was found for the EK sample. (See Table 5.)

In the EF sample, where students have one less year of instruction, an interaction was found; $F_{8,1506} = 3.11$, p .002. This interaction explains only 2 percent of the variance. Students in the lowest IQ block (below 70) gain significantly less during third (but not second) grade than many of their peers. One interpretation of this phenomena is that the extra year of instruction in kindergarten is particularly important for children entering with the lowest skills. The high IQ children (111 and above) also gain significantly less than their peers during third grade. This finding is harder to interpret. With those two exceptions during third grade, the pattern of basically similar growth rates is again found. Less than 2 percent of the variance is explained by differential growth.

MAT Total Reading. A somewhat different picture emerges on MAT Total Reading (Figures 5 and 6, Table 8). All IQ blocks in both samples

---

Figures 5, 6, Table 8 about here

---

make less progress during third grade than second grade on Reading subtests of the MAT, with the exception of the very highest IQ block in the EK sample. This IQ block gains significantly more than the lowest IQ block, maintaining a score equivalent to the 87-88th percentile, whereas all other IQ blocks demonstrate losses comparable to the national norm sample. A comparable trend is noticed for the EF sample.

Whereas the first and second grade Reading tests use a limited vocabulary, the third grade (Elementary) version of the MAT uses a largely uncontrolled, adult vocabulary parralleling the practices of the basal reader series for the intermediate grades. Schools typically assume that children will build their vocabulary at home and offer virtually no formal instruction in word vocabulary or word knowledge. This does not appear to be the case for low income students; and the consistent drops in third grade reading for low income students not only in the Direct Instruction Model (Becker, 1977), but in virtually all Follow Through approaches (Stebbins et al., 1977, Vol. IVD) appear to reflect this problem. Note that this drop during the third grade does not occur in WRAT Reading *a measure of decoding) or MAT Math. Becker (1977) offers a thorough discussion of this issue.

The Slosson is an intelligence test with a heavy verbal loading, and it makes some sense that students entering with very high verbal skills would gain significantly more than their low IQ counterparts on the third grade reading measure with a heavy emphasis on vocabulary knowledge. During the second grade, however, all IQ blocks gain roughly the same amount, as evidenced by the parallel line segments.

Slosson Intelligence Test (SIT). Tables 10 and 11 present the results on the SIT. Entry scores for the highest and lowest IQ blocks are "corrected"

Tables 10 and 11

for regression artifacts using the Campbell and Erlebacher (1970) formula, though the authors would be the first to admit the correction is merely a crude estimate. Since students were placed in IQ blocks on the basis of initial SIT scores, one would expect regression phenomena. The host of significant interactions during the first year are therefore not surprising. The most surprising finding is that no significant interactions are found during the remaining years of the structured educational program.

As Nesselroade et al. (1980) discuss, in longitudinal analyses there are a variety of conceivable regression phenomena. In this case, there is evidence of regression towards the population mean during the first year and then, for the very high and low IQ blocks, some slight regression towards the "subpopulation" means during the remaining years. Note that the samples entering the program with IQ's over 111 do not lose during the Follow Through years though one might expect some repeated regression phenomena. The low IQ children, on the other hand, display appreciable gains, even after the entry IQ has been "corrected"--the below 71 blocks gains 17 points in the EK sample and 9.4 points in the EF sample; gains for the 71-90 block are 15.6 and 9.2, respectively. The Slosson was used as a test of general language competence, and it appears that after the first year, students are making roughly comparable growth.

Correlational Analyses

Tables 13 and 15 present concurrent and predictive validity correlation coefficients for EK and EF samples, respectively. These are accompanied by the descriptive statistics for each sample. As indicated previously, participation in this sample was limited to students present for all testing in

---

Tables 13, 14, 15, 16

---

the first and last year of the program. The correlations between IQ and achievement scores at the end of third grade are consistently the highest, for both students at sites that begin Follow Through in kindergarten (EK) (see Table 13) and for students at sites that begin Follow Through in first grade (see Table 15). The range of concurrent validity coefficients from .48 to .65 for EK children and .42 to .66 for EF children. These are in the range one would predict from Crano et al. (1972) and other studies. As expected the magnitude of the correlations is greater for those domains requiring high order cognitive processes such as MAT Reading Comprehension, Word Knowledge, Math Concepts, and Math Problem Solving, than those involving more mechanical skills--WRAT Reading, MAT Math Computation, and Language. (Note: MAT Language tests only the grammatical aspects of language.)

The difference between the concurrent validity of IQ and the predictive validity of IQ seems somewhat consistent, both within the EK sample (Table 13) and within the EF sample (Table 15). For example, the respective concurrent and pre-first grade predictive validities of IQ for the EF sample are .63 and .48 for MAT Total Math, .53 and .35 for Math Computation, and .60 and .47 for Math Concepts. No within-sample pattern emerges for differences between concurrent and predictive IQ validities, whether scores are grouped by academic area (Reading, Math, Language) or by cognitive level (basic skill scores like WRAT Reading, MAT Language, MAT Math Computation, higher cognitive scores like MAT subtests in Word Knowledge, Reading Comprehension, Problem Solving, Math Concepts).

However, the difference between concurrent and predictive validity is greater on every third grade score for the EK sample than the EF sample. This is due to the substantially lower predictive validities of IQ for the EK sample. This finding is consistent with the IQ literature that finds IQ scores before

the age of six years to be generally less indicative of future achievement than later IQ scores (McCall et al., 1973). Even the end-of-kindergarten IQ scores of the EK sample correlate less with third grade scores than the start-of-first grade IQs of the EF sample.

The best predictor of third grade achievement scores of all the measures used, for both EK and EF students, is the WRAT Reading score attained after one year of Follow Through. (See the end of K and end of 1 columns of Tables 13 and 16, respectively.) For the EK sample, the IQ upon entry predicts slightly better than the measured IQ after one year of instruction. However, it is the WRAT Reading after one year of instruction--the only achievement measure--that consistently registers the highest predictive validity coefficients, ranging from .33 to .46 with a median of .42). Note that the WRAT score predicts both scores in higher order operations and more mechanical basic skills. The IQ scores administered upon entry are relatively weak predictors, ranging from .21 to .41 with a median of .33.

For the EF sample, both beginning and end-of-first grade IQ predict equally well, though there are some minor non-systematic fluctuations in WRAT Reading, MAT Math Computation and Language. Again, the achievement measure given at the end of first grade, WRAT Reading, is the best predictor (range of .40 to .55, median of .51).

Correlations for Ethnic Subpopulations. Since the EF sample consisted of a large proportion of students for whom English is a second language, and prediction of IQ and achievement is a more questionable issue for these students (cf. Gandara, Keogh, & Yoshioka-Maxwell, 1980), the EF sample was broken down into three subpopulations--Anglo (Tables 16 and 17), English as a Second Language (ESL) (Tables 18 and 19), and Black (Tables 20 and 21). The decision

to separately analyze the Black and Anglo samples was stimulated by the controversies concerning predictive validity of IQ and achievement measures for nonwhite students (cf. Oakland, 1978l Boussard et al., 1980; Crano et al., 1972).

For the ESL students, WRAT Reading at the end of first grade is the best overall predictor; entry IQ is also a surprisingly strong predictor (median of .55). Unlike the EK sample, predictive validity is usually higher than concurrent valicity (median of .59). For the primarily rural Black sample, achievement and IQ predict all measures equally well (except the relatively mechanical skills involved in MAT Language and WRAT Reading). No systematic patterns can be discerned in the rural Anglo sample, except that, unlike the EK sample, predictive and concurrent validity coefficients for IQ are generally comparable, with a median value of .52 and .62, respectively.

Summary and Discussion of Correlational Analyses. In general, for the EK sample, entry IQ is a moderately weak predictor of third grade achievement. Concurrent validity coefficients, on the other hand, are of the general magnitude of .6, the typical value found in the elementary grades. The achievement measure administered after one year of instruction is a consistently stronger predictor. This would tend to support the learning potential conceptualization of Budoff et al. (1976)--that assessments of minority and handicapped children should be based on ability to acquire new learning (i.e., WRAT Reading at the end of kindergarten) rather than on tests of prior knowledge (i.e., the Slosson administered during the first few weeks of school).

However, a very different picture emerges for the EF sample. If one ignores WRAT Reading and MAT Language, both tests of relatively "lower order" cognitive processes, then entry IQ is as good a predictor as

achievement tests or the IQ test administered at the end of first grade. The general magnitude of the correlation is between .5 and .6 for Black, rural Anglo and Hispanic samples. This is in almost the same range as those found by Crano et al. (1972). The magnitude of the concurrent and predictive correlations for IQ are, as Crano et al. found, roughly the same. Thus, in a classical sense, IQ is a moderately good predictor, explaining 30 to 35 percent of the variance in achievement three and a half years later.

Yet, it is equally important to realize that 65 percent of the variance in achievement is not explained by IQ. And that, as our longitudinal analyses of variance demonstrate in a controlled setting, IQ does not predict academic learning rate in classrooms to a meaningful degree.

Summary and Conclusions

The present study explored the relationship between entry IQ and yearly academic growth rates on the Wide Range Achievement Test, the Metropolitan Achievement Test, and the Slosson Intelligence Test for low income students in 20 communities involved in the Direct Instruction Follow Through Project. The curriculum materials used, the teaching techniques used, and the allocation of time and personnel in the classroom were all clearly specified and monitored. (See Becker & Carnine. 1980; Becker, Engelmann, Carnine & Maggs, 1981.) In this respect these analyses differ from previous work in this area (e.g., Crano et al, 1972; Oakland, 1978).

The major finding was that with this educational model, little, if any, relationship was found between entry IQ and yearly learning rate. Generally, the students entering the program at higher cognitive skill levels finished third grade at higher levels, but, as Anderson (1939) predicted, the growth rates (or amount of material learned) were not significantly different between IQ blocks. This included both basic

skills areas such as the word analysis skills tested in WRAT Reading and the problem solving and conceptual abilities measures by MAT Total Math and Total Reading.

There were a few minor exceptions to this picture: artifactual first year effects on the Slosson and third grade MAT Total Reading for the EK sample, first grade WRAT Reading and third grade MAT Total Reading for the EF sample; these are discussed above. The most striking finding is how few significant interactions were found. With the large sample sizes (of 600-1000) utilized in this analysis, one is assured that the power to detect true population differences is more than adequate, and the conclusions reliable. For each achievement test, less than 3 percent of the variance is explained by the interaction of IQ block and academic achievement. Thus, IQ does not predict yearly learning rate, or ability to gain from instruction.

Correlational analyses indicate moderately high concurrent validity coefficients between IQ scores and achievement test scores. For the EF sample, IQ scores at entry (beginning of first grade) correlate about as well with end of third grade achievement scores as third grade IQ correlates with third grade achievement. This parallels Crano et al's finding. The magnitude of the correlations is in the moderate range (.5 to .6), again replicating previous work. The same was not true for the EK sample; here, entry IQ was not a strong predictor of third grade achievement. WRAT Reading scores administered after one year of instruction were superior predictors of third grade achievement.

The most consistent finding is that though IQ scores demonstrate moderately concurrent and, for the EF sample, predictive validity, they do not predict ability to learn or profit from instruction.

These patterns of finding would be of only academic interest if this were not an effective educational model for low income students. Mean per-

formance at the end of third grade corresponds to roughly the 42nd percentile in MAT Total Reading and the 58th percentile in MAT Total Math for the EK sample, and the 38th and 59th percentiles, respectively, for the EF sample. For low income students in compensatory education programs, these are well above expected values, which tend to be between the 20th and 30th percentiles for EK children, and somewhat lower for the EF communities (cf. Molitor et al., 1977; Stebbins et al., 1977; Ozenne et al., 1975). These secondary analyses document the effectiveness of this educational model for the low performing students as well as the middle-level and high achieving groups.

It is obviously important to see to what extent these patterns are found in other compensatory educational models, including the traditional basal approaches used in the majority of schools. The longitudinal analyses of variance offer information about growth rates that supplement the traditional correlational approaches typically used (cf. Bossard et al, 1980; Crano et al., 1972; Oakland, 1978). But several questions of real importance for both compensatory education and special education immediately come to mind about the current analyses.

There appear to be some feature(s) of this model that allow for strong growth patterns for low performers without hurting the high performing students. Yet in a field experiment like Follow Through, it is impossible to isolate each variable from the composite. All the authors can do at this point is offer some tenable hypotheses/conjectures about features of the Direct Instruction model that may help explain this phenomenon.

The Direct Instruction approach is <u>not</u> individualized in the sense of, say, the Individualized Education Follow Through model (Leinhardt, 1977); yet there is a clear concern with each student beginning all academic programs at a level appropriate for his or her current skill level. By breaking

each class into 3 or 4 instructional groups for Reading and Arithmetic, this can almost always be achieved. To do so, in most instances, paraprofessional aides must do some of the teaching--by using scripted lessons and providing inservice training, this, too, is quite feasible. Thus, though the actual class size may be 25 or 28, for purposes of instruction in the core areas of reading, language and arithmetic, the "effective class size" is six to ten. Glass (1979) has found that more learning takes place in classes of 15 or less. The smaller classes allow for the intense teacher-student interactions and quantities of direct feedback from teachers that Rosenshine and Berliner (1978) and Fischer et al. (1980) have found to be consistently correlated with high academic gains in the primary grades. Small group instruction allows for the feedback system necessary for mastery learning (Bloom, 1976) rather than the "teach down the middle" approach often used in traditional classrooms. "Teaching down the middle" would seem particularly disastrous for low performers.

The Direct Instruction model has always been a mainstreamed model. Even students entering with measured IQs below 70 remain in the classroom for reading and arithmetic and language instruction rather than leaving the classroom for special education or remedial services. A key issue in inservice training is working with teachers on methods for teaching low performing groups. By not pulling these students out of the classroom, they are spared possible stigmatization. Furthermore, the classroom teacher is made to feel responsible for these students' progress. Cooley (1980) and Cazden, Green, and Rosen (1980) have indicated that the diffusion of responsibility typical of "pullout" programs may have a negative impact on low performing children.

There are several factors to this model that may contribute to its success with both low and high performing students--the use of curriculum materials that begin at a low level and use task analysis to break down

each cognitive task into its component parts, mastery learning, continuous assessment of student progress through criterion referenced tests, the allocation of time during the school day to allow for a strong academic emphasis, and a mainstreaming rather than a "pullout" model for low performing children. It is unclear which of these components, or what combination of these components, is most strongly related to the success with both low and high performers. Nor do we have comparable information on yearly academic growth rates for low income students in the primary grades with other educational models.

One thing is clear, though. With this educational model, even though in the traditional sense IQ scores at the age of six demonstrate moderate predictive validity coefficients, IQ does little to predict yearly academic gains or ability to learn.

Anderson, J. E. The limitations of infant and preschool tests in the measurement of intelligence. Journal of Psychology, 1939, 8, 351-379.

Becker, W. C., & Carnine, D. Direct instruction: An effective approach for educational intervention with the disadvantaged and low performers. In B. J. Lahey & A. E. Kazdin (Eds.), Advances in clinical child psychology. New York: Plenum Press, 1980.

Becker, W. C., & Engelmann, S. Analysis of achievement data on six cohorts of low income children from 20 school districts in the University of Oregon Direct Instruction Follow Through Model. Follow Through Project, Technical Report #78-1. Eugene, OR: University of Oregon, 1978.

Becker, W.C., & Engelmann, S. Teaching 3: Evaluation of instruction. Chicago: Science Research Associates, 1976.

Becker, W. C., Engelmann, S., Carnine, D., & Maggs, A. Direct instruction technology--Making learning happen. In P. Karoly & J. J. Steffen (Eds), Advances in Child Behavior Analysis and Therapy: Vol 2. New York: Gardner Press, 1981.

Bereiter, C., & Kurland, M. A constructive look at Follow Through results. Interchange, in press.

Bloom, B. Human characteristics and school learning. New York: McGraw-Hill, 1976.

Bossard, M. D., Reynolds, C. P., & Gutkin, T. B. A regression analysis of test bias on the Stanford -Binet Intelligence Scale for Black and White children referred for psychological services. Journal of Clinical Child Psychology, 1980, 9, 52-54.

Budoff, M., & Corman, L. Effectiveness of a learning potential procedure in improving problem-solving skills of retarded and nonretarded children. American Journal of Mental Deficiency, 1976, 81(3), 260-64.

Budoff, M., & Hmailton, J. Optimizing test performance of moderately and severely mentally retarded adolescents and adults. American Journal of Mental Deficiency, 1976, 81, 49-57.

Campbell, D. T., & Erlebacher, A. How regression artifacts in quasi-experimental evaluations in compensatory education tend to underestimate effects. In J. Hellmuth (Ed.), Disadvantaged child: Compensatory education--A national debate (Vol. 3). New York: Brunner/Mazel, 1970.

Crano, W. D., Kenny, D. A., & Campbell, D. T. Does intelligence cause achievement? A cross-lagged panel analysis. Journal of Educational Psychology, 1972, 63, 258-275.

Cronbach, L. J. Appendix I: Measuring mental ability--Lingering questions and loose ends. In B. D. Davis & P. Flaherty (Eds.), Human diversity: Its causes and social significance. Cambridge, MA: Ballinger, 1976.

Cronbach, L. J., & Furby, L. How we should measure "change"--or should we? Psychological Bulletin, 1970, 74, 68-80.

Dixon, W. J., Ed. BMD biomedial computer programs. Los Angeles: University of California Press, 1977.

Fischer, C. W., Berliner, D. C., Filby, N. N., Marliave, R., Cahen, L. S., & Dishaw, M. M. Teaching behaviors, academic learning time and student achievement: An overview. In C. Denham & A. Lieberman (Eds.), Time to learn. Washington, D.C.: USOE/NIE Printing, 1980.

Friedman, H. Magnitude of experimental effect and a table for its rapid estimation. Psychological Bulletin, 1968, 70, 245-251.

Gandara, P., Keogh, B. K., & Yoshioka-Maxwell, B. Predicting academic performance of Anglo and Mexican-American kindergarten children. Psychology in the Schools, 1980, 17, 174-177.

Gersten, R. The search for the cognitive deficit in autism: Beyond the stimulus overselectivity model. The Journal of Special Education, 1980, 14, 47-65.

Glass, G. V., & Smith, M. L. Meta-analysis of research on class size and achievement. Educational Evaluation and Policy Analysis, 1979, 1, 2-16.

Gulliksen, H. Theory of mental tests. New York: Wiley, 1950.

House, E. R., Glass, G. V., McLean, L. D., & Walker, D. E. No simple answer: Critique of the "Follow Through" evaluation. Harvard Educational Review: 1978,

Jensen, A. R. Bias in mental testing. New York: The Free Press, 1980.

Leinhardt, G. Modeling and measuring educational treatment in evaluation. Review of Educational Research, 1980, 50, 393-420.

Leinhardt, G. Program evaluation: An empirical study of individualized instruction. American Educational Research Journal, 1977, 14, 277-293.

Linn, R. L., & Slinde, J. A. The determination of the significance of change between pre- and post-testing periods. Review of Educational Research, 1977, 47, 121-150.

McCall, R. B., Appelbaum, M. I., & Hogarty, P. S. Developmental changes in mental performance. Monographs of the Society for Research in Child Development, 1973, 38(3, Serial No. 150).

Meyers, L. Comparison of learning rate and retention between a mongoloid and normal population. In W. C. Becker & S. Engelmann, Teaching 3: Evaluation of instruction. Chicago: SRA, 1976.

Molitor, J., Watkin, N., Napior, D., & Proper, E. C. Education as experimentation--The non-Follow Through study. Cambridge, Mass.: Abt Associates, 1977.

Nesselroade, J. R., Stigler, S. M., & Baltes, P. B. Regression toward the mean and the study of change. Psychological Bulletin, 1980, 88, 622-637.

Oakland, T. Predictive validity of readiness tests for middle and lower socio-economic status Anglo, Black, and Mexican American children. Journal of Educational Psychology, 1978, 4, 574-582.

Olson, G. M., et al. Long-term correlates of children's learning and problem-solving behavior. Journal of Educational Psychology, 1968, 59, 227-232.

Ozenne, D., et al. United States Office of Education. Annual evauation report on programs administered by the U.S. Office of Education, FY 1975. Washington, D.C.: Capital Publications, Educational Resource Division, 1976.

Rosenshine, B. V., & Berliner, D. C. Academic engaged time. British Journal of Teacher Education, 1978, 4, 3-16.

Stebbins, L., St. Pierre, R. G., Proper, E. C., Anderson, R. B., & Cerva, T. R. Education as experimentation: A planned variation model (Vol. IVA-D). Cambridge, MA: Abt Associates, 1977.

Stevens, S. S. (Ed.) Handbook of experimental psychology. New York: Wiley, 1951.

Stevenson, H. W., & Odom, R. D. Interrelationships in children's learning. Child Development, 1965, 36, 7-19.

Thorndike, R. L. Concepts of culture-fairness. Journal of Educational Measurement, 1971, 8, 63-70.

Vane, J. R. Relation of early school achievement to high school achievement when race, intelligence, and socioeconomic factors are equated. Psychology in the Schools, 1966, 3, 124-129.

Wilhelm, H., & Lovaas, O. I. Stimulus overselectivity: A common feature in autism and mental retardation. American Journal of Mental Deficiency, 1976, 81, 227-241.

Winer, B. J. Statistical principles in experimental design, 2nd ed. New York: McGraw-Hill, 1971.

Wolf, R. Review of Metropolitan Achievement Test. In O. K. Buros (Ed.), The sixth mental measurements yearbook. Highland Park, NJ: The Gryphon Press, 1978.

Zeaman, D. One programmatic approach to retardation. In D. K. Routh (Ed.) The experimental psychology of mental retardation. Chicago: Aldine, 1973.

Table 1

Coefficients for Predictive Validity of Slosson Intelligence Test for Anglo, Mexican-American and Black Children-from Oakland (1978)

				Grade 2 Metropolitan Achievement Tests		Grade 4 California Achievement Tests			Grade 2 California Test of Mental Maturity		
				Reading	Arith		Reading	Math	Lang	Non-Lan	Total
Group	M	SD	N	Total	metic	N	total	total	IQ	guage IQ	IQ
Total				.32**	.42**		.66**	.52**	.70**	.67**	.73**
Total Anglo	104.7	17.4	86	.46*	.71**	66	.67**	.27	.70**	.74**	.72**
Middle class	114.9	14.3	43			31	.59*	.05	.60**	.48	.65*
Lower class	93.9	13.6	43	.04	.70*	35	.71*	.54	.59*		
Total Mexican American	91.2	15.8	106	.60**	.44*	95	.76**	.68**	.76**	.75**	.80**
Middle Class	98.3	13.3	41	.59	.21	36	.79**	.59*	.92**	.61*	.89**
Lower class	84.4	15.1	65	.56*	.35	59	.66**	.59*	.56*	.78**	.69**
Total Black	96.1	13.8	117	-.07	.17	90	.12	.19	.65*	.43*	.57**
Middle Class	101.4	15.3	52			40	-.07	-.08	.75**	.56*	.70**
Lower Class	90.9	9.5	65	.32	.59**	50	.26	.35	.69**	.41	.61*

* $p < .05$

** $p < .01$

Table 2

Primary Ethnic Compositions of Follow Through Sites

Type of Program	Primary Ethnic Composition	Sites
Entering kindergarten (EK) Sites (4 year programs)	Black	New York, NY (P.S. 137); Grand Rapids, MI; Chicago, IL; Dayton, UH; E. St. Louis, IL; Flint, MI; Providence, RI; Racine, WI; Washington, D.C.
	Native American	Rosebud Tribe, SD; Todd County, SD; Cherokee, NC
	Anglo	Flippin, MS; West Iron Co., MI
Entering First Grade (EF) Sites (3 year programs)	Black	Tupelo, MS; Williamsburg County, SC
	Hispanic	Uvalde, TX; Dimmit, TX; E. Las Vegas, NM
	Anglo	Smithville, TN

Table 3

Ethnic Description of Samples and Language Spoken in the Home

	Percent of Total Sample	
	Kindergarten Starting (EK)	First Starting(EF)
ETHNIC GROUP:		
Black	49	37
Caucasian	26	20
Indian	23	6
Mexican	*	21
Puerto Rican	*	*
Spanish	2	16
LANGUAGE SPOKEN AT HOME:		
English	88	62
Non-English	12	38

* Less than ½ percent.

Table 4. Testing Schedule for Five Cohorts of Follow Through Children (1969-77)

Sample	Test	Beginning of Kindergarten	End of Kindergarten	Beginning of Grade 1	End of Grade 1	End of Grade 2	End of Grade 3
Students from ing sites beginning in Kindergarten (EK)	Wide Range Achievement WRAT (level 1)	X	X		X	X	X
	Slosson Intelligence Test (SIT)	X	X		X	X	X
	Metropolitan Achievement Test[a] MAT				(Primary I level) X	(Primary II level) X	Elementary level) X
Students from sites beginning in grade 1 (EF)	WRAT			X	X	X	X
	SIT			X	X	X	X
	MAT[a]				(Primary I level) X	(Primary II level) X	Elementary level) X

[a] MAT testing began in 1972

Table 5

Full Longitudinal Testing Per Cohort

Sample	Cohort	Began	Ended	Full Longitudinal Testing On
Entering Kindergarten Students (EK)	I	Fall 1969	Spring 1973	SIT, WRAT
	II	Fall 1970	Spring 1974	WRAT
	III	Fall 1971	Spring 1975	WRAT, MAT
	IV	Fall 1972	Spring 1976	SIT, WRAT, MAT
	V	Fall 1973	Spring 1977	SIT, WRAT, MAT
Entering First Grade Students (EF)	I	Fall 1969	Spring 1972	SIT, WRAT
	II	Fall 1970	Spring 1973	SIT, WRAT
	III	Fall 1971	Spring 1974	WRAT
	IV	Fall 1972	Spring 1975	WRAT, MAT
	V	Fall 1973	Spring 1976	WRAT, MAT

Table 6. Summary of Analysis of Variance for WRAT Reading

Sample	F for Interaction	df	p	r_m^2 Proportion of Explained variance	Locus of Interaction (Scheffe Tests
Students from Entering-Kindergarten (EK) sites	0.82	20/2744	ns	<.01	None
Students from Entering-First (EF) Sites	4.61	12/4491	<.001	.01	During grade 1 - the highest block gained significantly more than the lowest block.

Table 7. Summary of Analysis of Variance for MAT Total Math

Sample	F	df	p	r_m^2 Proportion of explained variance	Locus of Interaction (Scheffe Tests)
Students from Entering-Kindergarten (EK) sites	1.54	10/2100	ns	<.01	None
Students from Entering-First (EF) sites	3.11	8/1506	<.002	.02	During grade 3, lowest and highest blocks gained significantly less than the other blocks.

Table 8. Summary of Analysis of Variance for MAT Total Reading

Sample	F for interaction	df	p	r_m^2 Proportion of Explained Variance	Locus of Interaction (Scheffe Tests)
Students from Entering-Kindergarten (EK) sites	5.15	10/2152	.001	.02	During grade 3, the highest block gained significantly more than the lowest block (which reflects the verbal content of the SIT and the Total Reading subtest of the Elementary level MAT).
Students from Entering-First (EF) sites	5.54	8/1510	.001	.03	During grade 3, the highest 2 blocks gained significantly more than the lowest block, and the highest block also gained significantly more than the second lowest block.

Table 9. Summary of Analysis of Variance for Slosson Intelligence Test

Sample	F for interaction	df	p	r_m^2 Proportion of Explained Variance	Locus of Interaction (Scheffe Tests)
Students from Entering-Kindergarten (EK) sites	12.82	20,2844	<.001	.08	During kindergarten, the lowest block gained significantly more than each of the other blocks; the highest blocks gained significantly less than each of the other blocks; block II gained significantly more than the 3 highest blocks. (There was no interaction for other grades. Results probably primarily due to regression artifacts.)
Students from Entering-first (EF) sites	27.51	12,4218	<.001	.07	During grade 1, the lowest block gained significantly more than each of the other blocks; the highest block gained significantly less than each of the other groups; block II gained significantly more than block IV. (There was no interaction for other grades. Results probably primarily due to regression artifact

Table 10

Slosson Intelligence Test for Kindergarten Starting(EK) Low Income Children*
N=717

Grade	IQ Block											
	70 and Below		71-90		91-100		101-110		111-130		131 and above	
	Mean IQ	S.D.	Mean IQ	S.D.	Mean IQ	S.D.	Mean IQ	S.D.	Mean IQ	S.D.	Mean IQ	S.D.
Entry-kindergarten	66.1**	6.2	82.8	5.1	95.9	2.9	105.6	2.8	117.5	4.9	133.1**	5.5
Post-kindergarten	92.7	9.0	101.3	12.4	110.7	11.7	117.1	11.7	123.3	13.3	134.0	11.9
Post-1	94.6	11.7	99.4	10.4	106.8	10.7	113.0	11.4	120.2	11.8	131.2	11.5
Post-2	88.1	12.0	97.9	13.3	105.8	13.7	112.2	13.4	118.0	15.5	131.0	18.4
Post-3	83.8	11.3	98.4	13.2	103.9	12.5	108.6	13.5	116.5	14.5	132.2	10.8
N	15		158		220		246		169		24	

* Significant interactions found during kindergarten for
(a) below 70 IQ block and all other IQ blocks
(b) 70-90 IQ block and 3 highest IQ block
(c) 90-100 IQ block and 2 highest IQ blocks
(d) 131+ IQ block and all other IQ blocks

** Means adjusted for regression effects.

Table 11

Slosson Intelligence Test for First-Starting(EF) Low Income Children*
N = 1086

IQ Block

Grade	Less Than 71		71-90		91-100		101-110		More than 110	
	Mean IQ	S.D.	Mean IQ	S.D.	Mean IQ	S.D.	Mean IQ	S.D.	Mean IQ	S.D.
Entry First	69.9**	4.4	83.0	5.3	95.4	2.9	104.8	2.8	116.6**	6.9
Post-First	81.1	11.3	91.5	10.6	101.6	10.6	106.1	10.4	110.8	13.2
Post 2	80.8	10.7	91.8	10.9	101.5	11.5	109.0	12.1	113.1	15.3
Post 3	79.3	12.1	92.2	14.0	102.4	13.6	111.5	14.0	115.5	16.1
N	43		409		327		175		132	

* Significant interactions during first grade for
(a) below 70 IQ block and all other IQ blocks
(b) 70-90 IQ block and 2 highest IQ blocks
(c) 90-100 IQ block and 2 highest IQ blocks
(d) 100-110 IQ block and 111+ IQ block

** Means adjusted for regression effects.

Table 12. Descriptive Statistics for All EK Students (n = 594)

	Mean Scale Score	SD	%ile Equiv.
IQ			
Pre-K	102.8	14.2	57th
Post 3rd	107.2	14.7	67th
Reading--Post 3rd			
WRAT Rdg.	5.05[a]	1.9	(est) 90th
MAT Total Rdg.	58.4	10.1	41st
MAT Word Knowledge	60.4	9.1	41st
MAT Reading Comp.	58.1	12.2	40th
Math--Post 3rd			
MAT Total Math	71.8	11.3	55th
MAT Math Comput.	71.8	12.2	73rd
MAT Prob. Solv.	66.5	11.6	46th
MAT Math Concepts	66.2	11.3	47th

[a] GE

Table 13. Correlations for EK Students (n=594)

Third Grade Scores	Predictive (start of K)		Predictive (end of K)		Concurrent
	Pre-K IQ	Pre-K WRAT Rdg.	Post K IQ	Post K WRAT Rdg.	Post 3rd IQ
IQ	.52	.34	.51	.43	1.00
READING					
WRAT Rdg.	.27	.26	.27	.42	.55
MAT Total Rdg.	.41	.36	.35	.43	.64
MAT Word Know.	.40	.35	.35	.46	.64
MAT Rdg. Comp	.36	.33	.28	.33	.57
MATH					
MAT Total Math	.34	.33	.29	.40	.63
MAT Math Comp.	.25	.26	.21	.33	.47
MAT Prob. Solv.	.33	.31	.27	.38	.59
MAT Math Conc.	.35	.35	.34	.42	.65
LANGUAGE					
MAT Language	.21	.26	.20	.45	.48

Table 14. Descriptive Statistics for All EF Students (n = 735)

	Mean Scale Score	SD	%ile Equiv.
IQ			
Pre 1st	92.8	12.5	33rd
Post 3rd	100.9	16.1	52nd
Reading--Post 3rd			
WRAT Rdg.	4.78[a]	1.70	(est) 68th
MAT Total Rdg.	57.6	8.9	38th
MAT Word Knowledge	59.4	8.5	38th
MAT Rdg. Comp.	58.3	9.9	41st
Math--Post 3rd			
MAT Total Math	72.9	10.4	60th
MAT Math Comp.	73.4	11.5	77th
MAT Prob. Solv.	67.4	10.7	49th
MAT Math Concepts	66.6	10.6	48th

Table 15. Correlations for EF Students
(n=735)

Third Grade Scores	Predictive (start of 1)		Predictive (end of 1)		Concurrent
	Pre 1st IQ	Pre 1st WRAT Rdg	Post 1st IQ	Post 1st WRAT Rdg	Post 3rd IQ
IQ	.61	.39	.64	.40	1.00
READING					
WRAT Rdg	.37	.39	.28	.53	.42
MAT Total Rdg	.57	.54	.54	.54	.61
MAT Word Know	.54	.53	.50	.52	.66
MAT Rdg Comp	.53	.47	.52	.51	.58
MATH					
MAT Total Math	.48	.37	.50	.49	.63
MAT Math Comp	.35	.28	.42	.40	.53
MAT Prob Solv	.48	.35	.47	.43	.59
MAT Math Conc	.47	.37	.47	.49	.60
LANGUAGE					
MAT Language	.37	.41	.32	.55	.46

Table 16. Descriptive Statistics for Anglo White Students (n = 218)

	Mean Scale Score	SD	%ile Equiv.
IQ			
Pre 1st	91.4	12.8	30th
Post 3rd	102.7	17.0	56th
Reading--Post 3rd			
WRAT Rdg.	4.68[a]	1.36	(est) 66th
MAT Total Rdg.	58.5	8.9	41st
MAT Word Know.	59.8	8.6	39th
MAT Rdg. Comp.	59.6	10.0	44th
Math--Post 3rd			
MAT Total Math	75.7	10.0	67th
MAT Math Comp.	77.9	10.9	88th
MAT Prob. Solv.	69.5	11.4	54th
MAT Math Concepts	68.4	10.1	57th

[a] GE

Table 17. Correlations for Rural Anglo Students (n=218)

Third Grade Scores	Predictive (start of 1)		Predictive (end of 1)		Concurrent
	Pre 1 IQ	Pre 1 WRAT Rdg	Post 1 IQ	Post 1 WRAT Rdg	Post 3rd IQ
IQ	.67	.37	.69	.38	1.00
READING					
WRAT Rdg.	.35	.43	.29	.59	.37
MAT Total Rdg	.64	.62	.61	.54	.65
MAT Word Know.	.57	.63	.55	.51	.57
MAT Rdg. Comp.	.64	.51	.63	.51	.66
MATH					
MAT Total Math	.56	.36	.57	.40	.70
MAT Math Comp.	.40	.27	.44	.33	.58
MAT Prob. Solv.	.52	.33	.53	.46	.62
MAT Math Conc.	.57	.36	.56	.46	.68
LANGUAGE					
MAT Lang.	.38	.41	.37	.54	.50

Table 18. Descriptive Statistics for ESL Students (n = 87)

	Mean Scale Score	SD	%ile Equiv.
IQ			
Pre 1st	84.8	11.8	18th
Post 3rd	96.1	13.6	40th
Reading--Post 3rd			
WRAT Rdg.	4.64[a]	1.27	(est) 64th
MAT Total Rdg.	55.4	7.0	32nd
MAT Word Knowledge	57.4	5.9	31st
MAT Rdg. Comp.	56.0	8.6	34th
Math--Post 3rd			
MAT Total Math	72.1	9.4	56th
MAT Math Comp.	73.1	9.6	76th
MAT Prob. Solv.	66.9	9.2	48th
MAT Math Concepts	64.8	10.4	44th

[a] GE

Table 19. Correlations for ESL Students (n=87)

Third Grade Scores	Predictive (start of 1)		Predictive (end of 1)		Concurrent
	Pre 1st IQ	Pre 1st WRAT Rdg	Post 1st IQ	Post 1st WRAT Rdg	Post 3rd IQ
IQ	.60	.41	.60	.54	1.00
READING					
WRAT Rdg	.40	.37	.11	.65	.46
MAT Total Rdg	.59	.44	.33	.59	.51
MAT Word Know	.60	.44	.34	.60	.54
MAT Rdg Comp	.55	.43	.33	.59	.45
MATH					
MAT Total Math	.61	.38	.48	.64	.57
MAT Math Comp	.55	.35	.47	.57	.50
MAT Prob Solv	.59	.30	.35	.51	.48
MAT Math Conc	.54	.42	.46	.55	.58
LANGUAGE					
MAT Language	.45	.43	.26	.72	.41

Table 20. Descriptive Statistics for Black Students (n = 368)

	Mean Scale Score	SD	%ile Equiv.
IQ			
Pre 1st	94.5	11.9	37th
Post 3rd	99.6	16.0	49th
Reading--Post 3rd			
WRAT Rdg.	4.83[a]	1.98	(est)69th
MAT Total Rdg.	56.8	9.0	36th
MAT Word Knowledge	58.9	8.9	36th
MAT Rdg. Comp.	57.2	9.8	38th
Math--Post 3rd			
MAT Total Math	70.5	10.5	50th
MAT Math Comp.	69.9	11.4	68th
MAT Prob. Solv.	65.8	10.8	43rd
MAT Math Concepts	64.8	10.7	44th

[a] GE

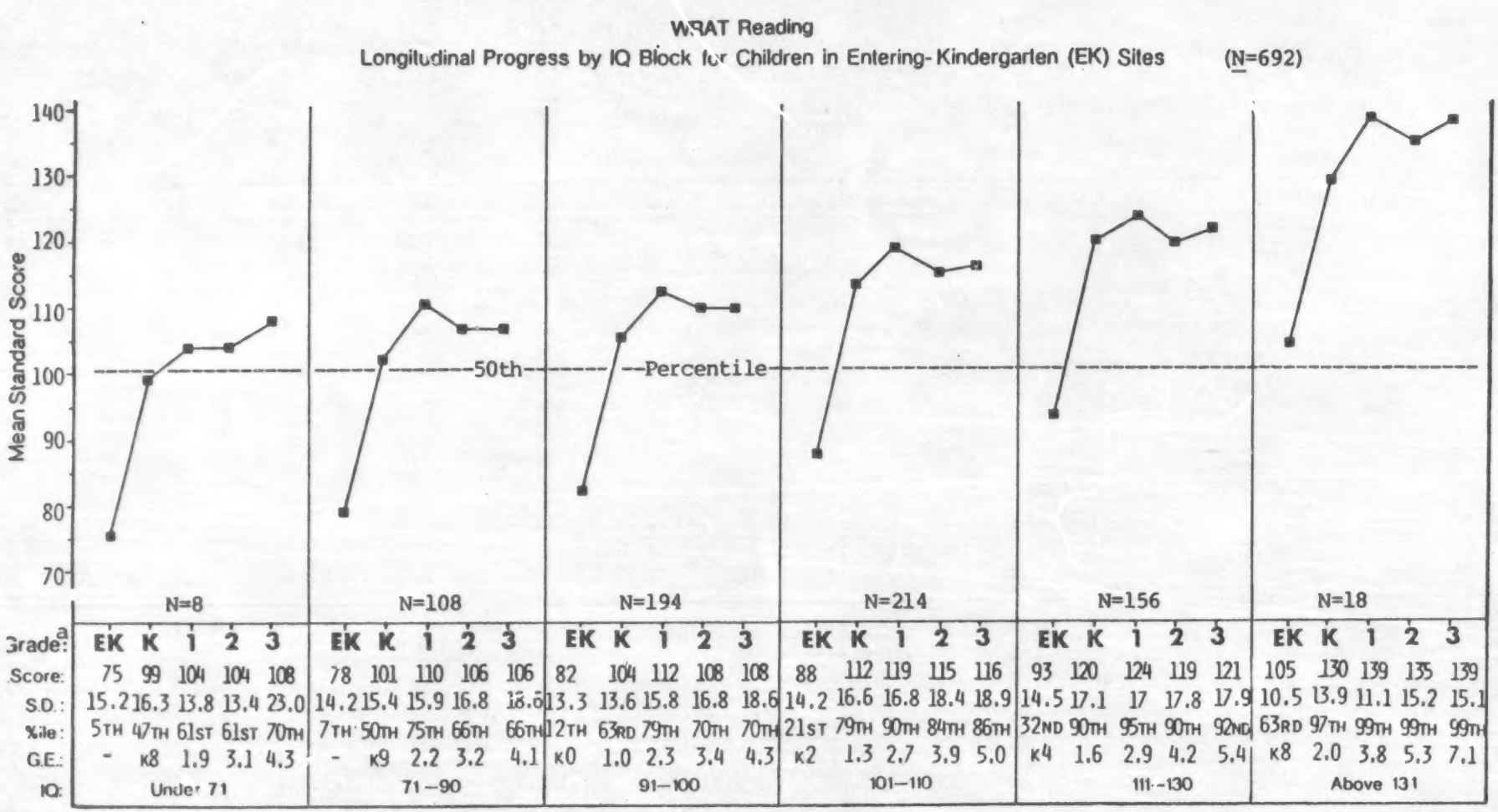
Table 21. Correlations for Black Students (n=368)

Third Grade Scores	Predictive (start of 1)		Predictive (end of 1)		Concurrent
	Pre 1st IQ	Pre 1st KRAT Rdg.	Post 1st IQ	Post 1st WRAT Rdg	Post 3rd IQ
IQ	.58	.42	.60	.37	1.00
READING					
WRAT Rdg	.37	.40	.29	.53	.45
MAT Total Rdg	.52	.53	.51	.52	.57
MAT Word Know	.51	.51	.49	.50	.56
MAT Rdg Comp	.47	.48	.47	.47	.52
MATH					
MAT Total Math	.48	.45	.45	.48	.58
MAT Math Comp	.38	.38	.38	.40	.48
MAT Prob Solv	.48	.41	.42	.46	.56
MAT Math Conc.	.43	.41	.40	.46	.52
LANGUAGE					
MAT Language	.33	.41	.32	.56	.44

Table 22. Intercorrelations Among Post 3rd Grade Tests for EK Students

	IQ	WRAT Rdg.	MAT Total Rdg.	MAT Word Know.	MAT Rdg. Compre.	MAT Lang.	MAT Total Math	MAT Prob. Solv.	MAT Math Concepts	MAT Math Computation
IQ	1.00	.55	.64	.63	.57	.48	.63	.59	.65	.48
WRAT Rdg.		1.00	.71	.70	.68	.75	.62	.56	.61	.48
MAT Total Rdg.			1.00	.95	.93	.70	.74	.69	.74	.58
MAT Word Know.				1.00	.80	.69	.70	.65	.72	.54
MAT Rdg. Compre.					1.00	.65	.73	.66	.69	.61
MAT Lang.						1.00	.62	.54	.59	.53
MAT Total Math							1.00	.91	.92	.89
MAT Prob. Solv.								1.00	.82	.69
MAT Math Concepts									1.00	.72
MAT Math Comput.										1.00

Figure 1

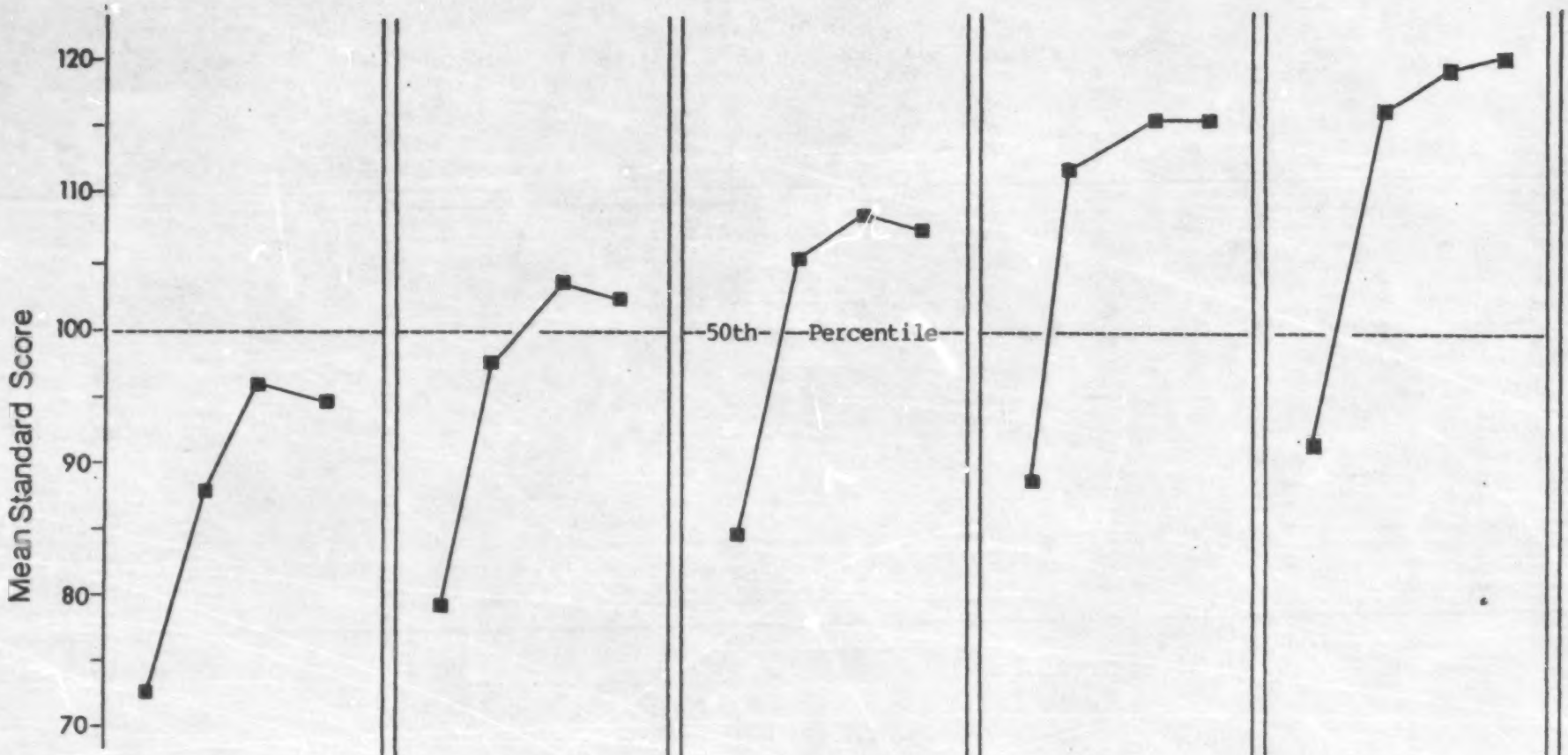


[a]All testing performed at end of academic year, except for EK

WRAT Reading

Longitudinal Progress by IQ Block for Children in Entering-First(EF) Sites[b]

(N=1,344)



	N=57				N=476				N=416				N=224				N=171			
Grade:[a]	EF	1	2	3	EF	1	2	3	EF	1	2	3	EF	1	2	3	EF	1	2	3
Id.Score:	72.3	87.4	95.8	94.4	78.9	96.9	103	102	84.4	105	108	107	88.9	111	115	115	90.7	115	119	120
S.D.:	10.3	11.7	14.8	11.9	9.0	14.8	14.9	14.2	10.7	15.1	16.3	14.5	11.6	17.9	16.8	17.6	12.7	19.5	20.3	20.6
%ile:	3rd	19th	39th	34th	8th	42nd	58th	55th	14th	63rd	70th	68th	23rd	77th	84th	84th	27th	84th	91st	91st
G.E.:	K1	1.0	2.5	3.0	K0	1.5	3.0	3.7	K2	1.9	3.4	4.2	K4	2.3	3.9	4.9	K5	2.5	4.3	5.4
IQ:	Under 71				71–90				91–100				101–110				Above 111			

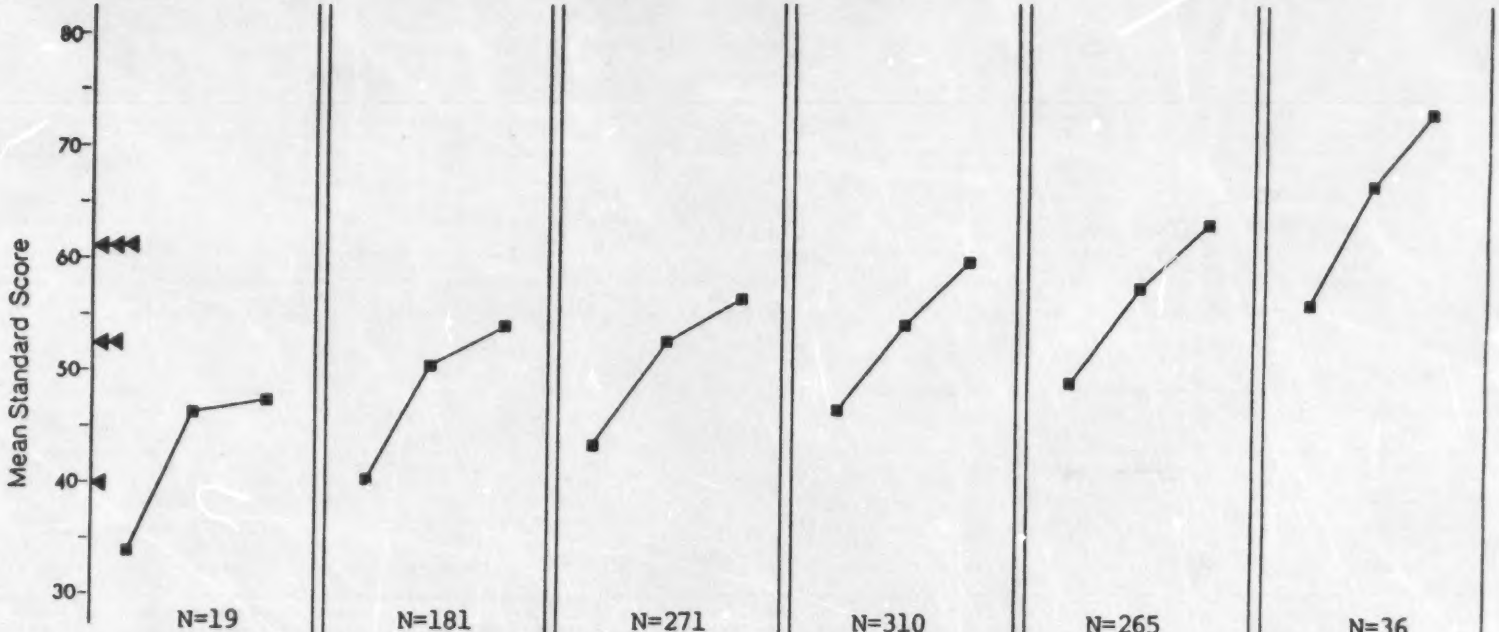
[a] All testing done at end of academic year, except for EF testing

[b] During grad. 1 the highest IQ Block gained significantly more than the lowest IQ Block

Figure 3

MAT Total Reading

Longitudinal Progress by IQ Block during 2nd and 3rd Grade for Children in Entering-Kindergarten (EK) Sites[b] (N=1,082)



Grade:[a]	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Std.Score:	33.9	46.6	47.4	40.1	50.6	54.3	43.1	52.7	56.0	46.0	54.5	59.6	49.0	57.1	63.2	55.4	65.1	72.6
S.D.:	7.7	7.0	13.1	8.2	6.5	8.4	8.6	7.5	8.6	8.4	8.1	8.5	9.3	8.7	9.6	8.1	9.2	7.7
%ile:	28TH	22ND	10TH	50TH	40TH	28TH	66TH	50TH	34TH	78TH	58TH	44TH	88TH	70TH	58TH	94TH	88TH	81ST
G.E.:	1.6	2.2	2.2	1.8	2.5	2.8	2.0	2.7	3.0	2.2	2.8	2.4	2.4	3.1	3.7	2.9	3.9	4.9
IQ:	Under 71			71–90			91–100			101–110			111–130			Above 131		

[a]All testing was performed at end of academic year

[b]During Grade 3, the highest IQ block gained significantly more than the lowest block (which reflects the verbal content of the SIT and the Total Reading subtest of the Elementary Level MAT)

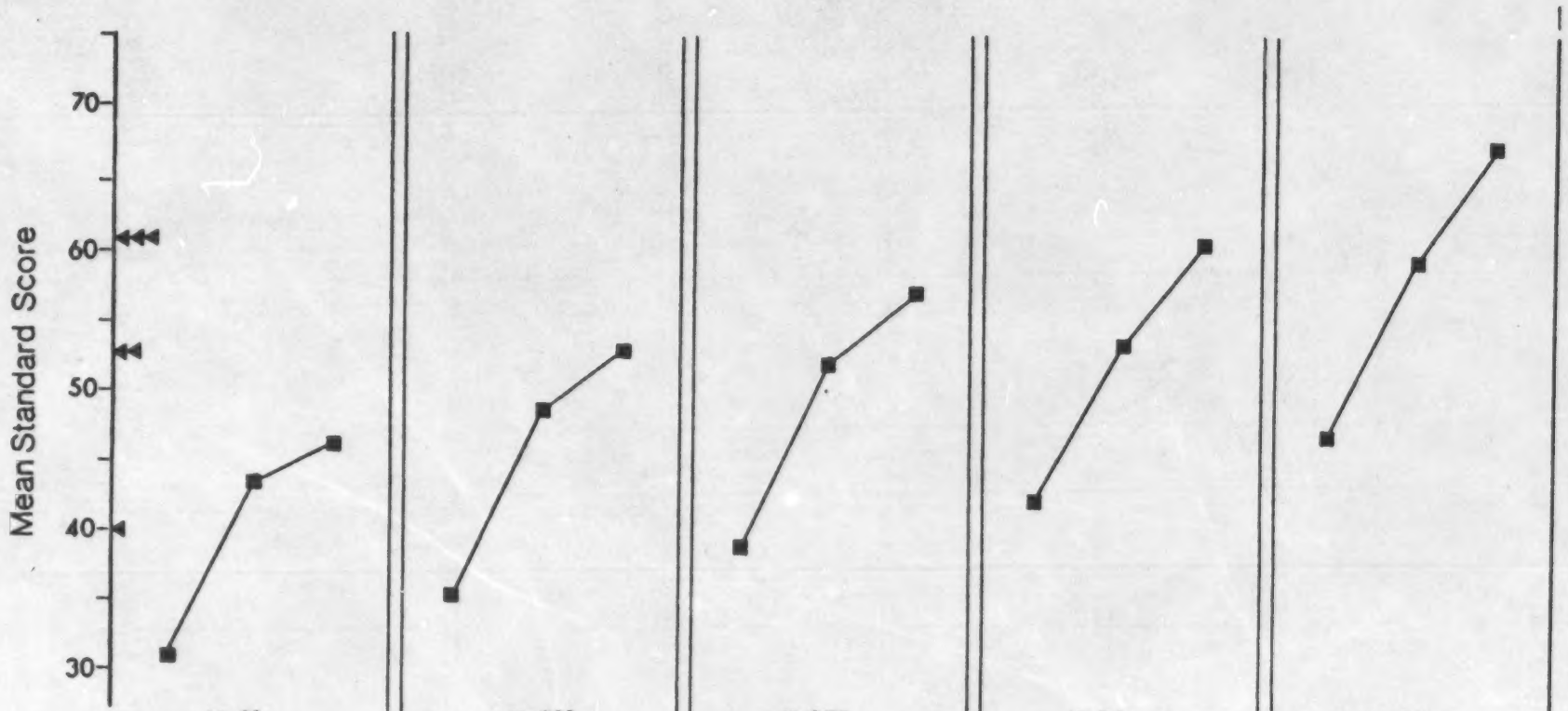
◀ Grade 1 National Median
◀◀ Grade 2 National Median
◀◀◀ Grade 3 National Median

Figure 4

MAT Total Reading

Longitudinal Progress by IQ Block during 2nd and 3rd Grade for Children in Entering-First(EF) Sites [b]

(N=521)



Grade:[a]	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Std.Score:	30.4	43.0	46.4	35.1	48.9	52.9	38.8	51.7	57.0	41.5	53.5	60.8	46.3	59.3	67.3
S.D.:	7.3	7.8	7.6	6.9	5.6	7.8	8.0	6.7	6.7	8.2	5.9	6.8	8.7	9.0	9.4
%ile:	20th	12th	9th	32nd	34th	24th	42nd	48th	36th	58th	52nd	50th	70th	74th	68th
G.E.:	1.5	2.0	2.2	1.6	2.4	2.7	1.8	2.6	3.1	1.9	2.7	3.5	2.2	3.3	4.2
IQ:	Under 70			70–90			91–100			101–110			111–130		

◀ Grade 1 National Median
◀◀ Grade 2 National Median
◀◀◀ Grade 3 National Median

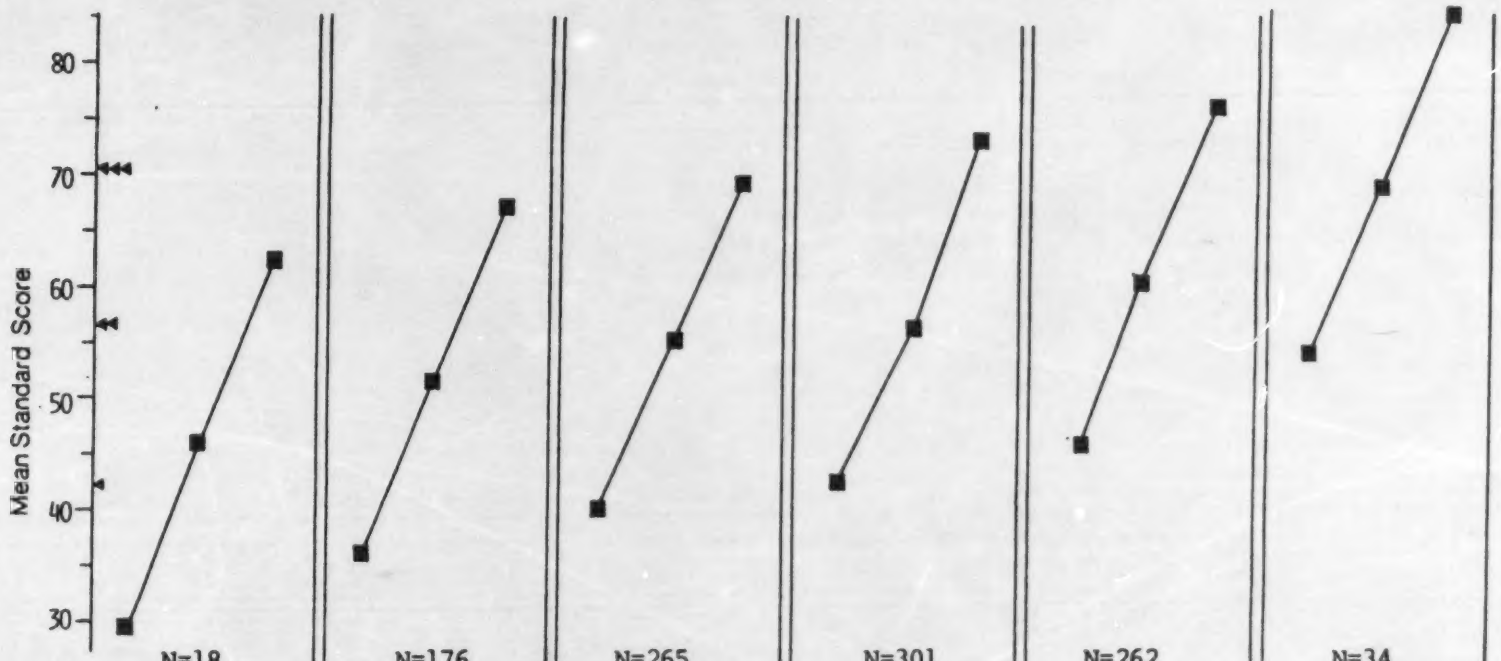
[a] All testing done at end of Academic year

[b] During grade 3, the highest 2 IQ blocks gained significantly more than the lowest block. The highest IQ block also gained significantly more than the

Figure 5

MAT Total Mathematics

Longitudinal Progress by IQ Block during 2nd and 3rd Grade for Children in Entering-Kindergarten (EK) Sites (N=1,056)



Grade[a]	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Std.Score:	29.2	46.1	62.6	35.8	51.4	67.7	40.0	55.2	69.7	42.6	56.9	73.5	45.6	60.2	76.4	54.4	69.2	85.1
S.D.:	7.3	10.1	9.5	8.0	8.5	10.6	9.0	8.2	9.7	9.2	8.8	10.7	10.8	10.0	11.6	10.6	8.5	8.1
%ile:	10TH	10TH	24TH	30TH	22ND	39TH	46TH	38TH	46TH	51ST	59TH	60TH	57TH	62ND	68TH	82ND	84TH	88TH
G.E.:	1.2	2.0	3.2	1.5	2.3	3.5	1.7	2.5	3.7	2.0	2.7	4.0	2.0	3.1	4.3	2.4	5.2	5.7
IQ:	Under 70			70–90			91–100			101–110			111–130			Above 131		

[a]All testing was done at the end of the academic year

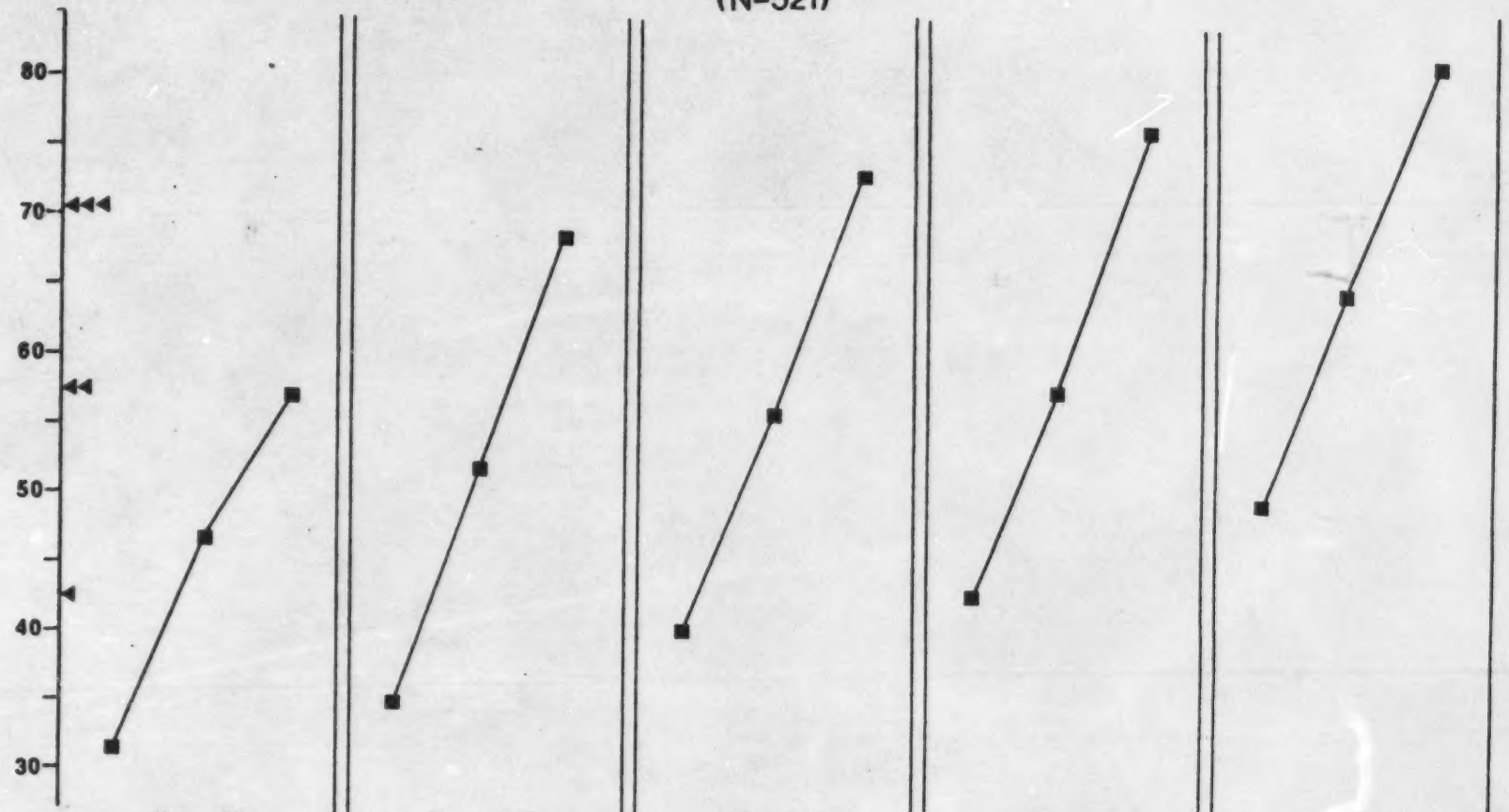
◀ Grade 1 National Median
◀◀ Grade 2 National Median
◀◀◀ Grade 3 National Median

Figure 6

MAT Total Mathematics

Longitudinal Progress by IQ Block during 2nd and 3rd Grade for Children in Entering-First(EF) Sites [b]

(N=521)



Grade: [a]	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Std.Score:	31.3	46.4	57.3	34.6	51.5	68.3	39.5	55.3	72.5	42.1	57.3	76.3	48.8	63.8	80.6
S.D.:	7.5	7.5	9.5	7.7	6.9	8.7	8.8	7.7	9.4	9.4	7.4	10.8	14.0	10.3	10.5
%ile:	14th	10th	12th	25th	22nd	40th	41st	38th	57th	50th	50th	68th	68th	77th	78th
G.E.:	1.2	2.0	2.7	1.5	2.3	3.6	1.7	2.5	3.9	1.8	2.7	4.3	2.2	3.4	4.8
IQ:	Under 71			70–90			91–100			101–110			111–130		

[a] All testing done at end of academic year

[b] During Grade 3 the lowest and highest IQ blocks gained significantly less than the other IQ blocks.

◂ Grade 1 National Median
◂◂ Grade 2 National Median
◂◂◂ Grade 3 National Median